Low-Beta/Vol Strategies and Their Role in Multi-factor Strategies

Zhe Wang

September 23, 2016

1 Introduction

Asset classes have long been the building blocks of investment portfolios, but when apparently uncorrelated investments moved in sync during the financial crisis, it raised fundamental questions about whether diversified portfolios actually were diversified.

Factor-based investing is one attempt to answer that question. By focusing on the underlying factors that define risk, return, and correlation this approach seeks to explain why some asset classes move together and to offer more efficient portfolio construction. Asset managers are starting to incorporate the idea into their portfolios, and a number of firms are offering factor-based mutual funds and ETFs.

Ever since 2009, when Ang, Goetzmann and Schaefer[1] produced their study for the Norwegian Government Pension Fund, many in the market have accepted the need to invest intentionally and efficiently in factor premia. A host of institutional mandates have evolved focusing on factors and, importantly, a series of financial products has come to market that has been built from these indices. Looking at the range of smart beta products in the market, many of which fall under the category of factor investing.

A factor can be thought of as any characteristic relating a group of securities that is important in explaining their return and risk. A large body of academic research highlights that long term equity portfolio performance can be explained by factors. This research has been prevalent for over 40 years; Barra for instance has undertaken the research of factors since the 1970s. Certain factors have historically earned a long-term risk premium and represent exposure to systematic sources of risk. Factor investing is the investment process that aims to harvest these risk premia through exposure to factors. We currently identify six equity risk premia factors: Value, Low Size, Low Beta/Volatility, Quality and Momentum. They are grounded in academic research and have solid explanations as to why they historically have provided a premium.

In this report, we will first study the low beta/volatility factor in detail, explain the economic rationale behind its excess return and describe how to create a index to catch its return. Then we will discuss how this factor could be integrated into a multi-factor basket.

2 Low-Beta(BAB)/Vol factor

The superior performance of low-volatility stocks - the low-volatility anomaly - has been documented to exist in equity markets around the globe. And since its discovery, a good amount of academic research has attempted to determine both its origins and whether or not it will continue to persist.

Among that research is a December 2013 paper[2], the authors of the study concluded that the reduction in a portfolios volatility is driven by a substantial decrease in its market beta, and that low-volatility strategies outperformed their corresponding cap-weighted market indexes due to exposure to the value factor, the betting-against-beta (BAB) factor as well as the duration factor.

In other words, investors were trading one risk (beta) for two others (value and term). This conclusion is consistent with the findings of prior research.

Here we will primarily focus on BAB factor. First, we give a sounding economic model explaining the rationale behind this factor. We then discuss how to build portfolios to catch this factor. Finally, we show our backtest results.

2.1 Limit on Leverages and its consequences

A basic premise of the capital asset pricing model (CAPM) is that all agents invest in the portfolio with the highest expected excess return per unit of risk (Sharpe ratio) and leverage or de-leverage this portfolio to suit their risk preferences.However,many investors,such as individuals,pension funds, and mutual funds, are constrained in the leverage that they can take and they therefore overweight risky securities instead of using leverage.

This behavior of tilting toward high-beta assets suggests that risky high-beta assets require lower risk-adjusted returns than low-beta assets, which require leverage. Indeed, the security market line for U.S. stocks is too flat relative to the CAPM [3] and is better explained by the CAPM with restricted borrowing than the standard CAPM[3, 4, 5], (see [6] for an excellent historical perspective).

Several questions arise: How can an unconstrained arbitrageur exploit this effect, i.e., how do you bet against beta? What is the magnitude of this anomaly relative to the size, value, and momentum effects? How does the return premium vary over time and in the cross section?

Frazzini and Pedersen explained these questions by considering a dynamic model of leverage constraints.

When the leveraged agents hit their margin constraint, they must deleverage. Therefore, the model predicts that, during times of tightening funding liquidity constraints, the low beta factor (or more precisely the BAB factor) realizes negative returns as its expected future return rises. Furthermore, the model predicts that the betas of securities in the cross section are compressed toward 1 when funding liquidity risk is high. Finally, the model implies that more-constrained investors overweight high-beta assets in their portfolios while less-constrained investors overweight low-beta assets and possibly apply leverage. Our model thus extends Black's[3] central insight by considering a broader set of constraints and deriving the dynamic time-series and cross-sectional properties arising from the equilibrium interaction between agents with different constraints.

2.2 Theory

We consider an overlapping-generations (OLG) economy in which agents $i = 1, ..., I$ are born each time period t with wealth W_t^i and live for two periods. Agents trade securities $s = 1, ..., S$, where security s pays dividends δ_t^s and has x^{*s} shares outstanding. Each time period t, young agents choose a portfolio of shares $x = (x^1, ..., x^S)$, investing the rest of their wealth at the risk-free return r^f , to maximize their utility

$$\max x^T(E_t(P_{t+1} + \delta_{t+1}) - (1 + r^f)P_t) - \frac{\gamma^i}{2} x^T \Omega_t x \tag{1}$$

where P_t is the vector of prices at time t, Ω_t is the variance-covariance matrix of $P_{t+1}+\delta_{t+1}$, and γ^i is agent i risk aversion. Agent i is subject to the following portfolio constraint:

$$m_t^i \sum_s x^s P_t^s \leq W_t^i \tag{2}$$

This constraint requires that some multiple m_t^i of the total dollars invested must be less than the agent i's wealth.

We are interested in the properties of the competitive equilibrium in which the total demand equals the supply:

$$\sum_i x^i = x^* \tag{3}$$

Consider the first order condition for agent i

$$0 = E_t(P_{t+1} + \delta_{t+1}) - (1 + r^f)P_t - \gamma^i \Omega x^i - \psi_t^i P_t \tag{4}$$

where ψ^i is the Lagrange multiplier of the portfolio constraint. Solving for x^i gives the optimal position.

The equilibrium condition now follows from summing over optimal positions

$$x^* = \frac{1}{\gamma}\Omega^{-1}(E_t(P_{t+1} + \delta_{t+1}) - (1 + r^f + \psi_t)P_t) \tag{5}$$

where the aggregate risk aversion γ is defined by $1/\gamma = \sum_i 1/\gamma^i$ and $\psi_t = \sum_i \frac{\gamma}{\gamma^i}\psi_t^i$ is the weighted average Lagrange multiplier.

Finally this leads to the equilibrium equation.

$$E_t[r_{t+1}^s] = r^f + \psi_t + \beta_t^s(E_t[r_{t+1}^m] - r^f - \psi_t) \tag{6}$$

where ψ_t is the average of some Lagrange multipliers which measures the tightness of leverage constraints.

This equilibrium equation immediately leads to the consequence that any portfolio with beta β_t^s will have an excess return α of CAPM sense

$$\alpha_t^s = \psi_t(1 - \beta_t^s) \tag{7}$$

One could notice that it is positive when $\beta_t^s < 1$ and negative when $\beta_t^s > 1$.

Therefore, we have shown theoretically that the BAB factor defined by $1 - \beta^i$ is a risk premium.

2.3 Portfolio construction

In order to show in reality that the BAB factor is indeed a systematic risk that brings excess return, one needs to construct a portfolio and back-tests it.

In their paper [7], Frazzini and Pedersen suggested a straightforward way to construct such a portfolio by explicitly choosing two baskets with highest x% betas and lowest x% betas. Then the portfolio simply short the high beta basket and long the low beta basket. So the portfolio has positive exposure to BAB factors in both long and short leg. We then neutralize the exposure to market by scaling the long and short legs to reach total beta neutrality.

Mathematically, noting the high and low beta basket as S^H and S^L, the initial weights w^i (in section 3 we will discuss different ways to determine initial weights)

Then, we could construct a beta-neutral portfolio by scaling the S^H and S^L so that their beta becomes -1 and 1 respectively. The final portfolio becomes

$$S = S^H \cup S^L, \quad \{w^i\}_{i \in S} = \left\{ \frac{w^i}{\beta^L} \right\}_{i \in S^L} \cup \left\{ -\frac{w^i}{\beta^H} \right\}_{i \in S^H} \tag{8}$$

So we could easily see that $\beta^S = 0$ and using (6) to see that

$$\mathbb{E}_t[r^S_{t+1}] = \frac{\beta^H_t - \beta^L_t}{\beta^H_t \beta^L_t} (\psi_t + r^f) > 0 \tag{9}$$

If we normalize the portfolio to be non-leveraged, then we go back to (7). Therefore, this is a simple method to construct beta-neutral portfolios that captures the excess returns from the BAB factor.

We also consider another method of constructing BAB factor driven portfolios. The motivation is from the observation that the above F-P method results in some large volatilities. Therefore, after allocating the initial weights, instead of neutralizing the portfolio to market, we scale the long and short leg to minimize the total risk.

$$\min_{\lambda} (w^L - \lambda w^H)^T \Omega (w^L - \lambda w^H) \tag{10}$$

The result is

$$\lambda = \frac{{w^H}^T \Omega_{LH} w^L}{{w^H}^T \Omega_{HH} w^H} \tag{11}$$

where Ω_{LH}, Ω_{HH} are the blocks of Ω corresponding to the basket S^H or S^L.

3 Implementation of the F-P method

In the implementation, the process is rather simple. The two baskets S^H and S^L are chosen simply to be the highest and lowest $x\%$ beta/vol stocks. The original weight $w^i, i \in S^k, k = L, H$ is determined by one of the two methods, the factor weighted and the rank weighted.

The factor weighted method is as following. One first identifies the two baskets S^H and S^L. Then one winsorizes the score vol) by shifting its mean to be 0 and capping all the factor value within 2.5 times standard deviation. Now one simply takes

$$w^i = -\hat{\beta}^i(\text{or} - \hat{\sigma}^i), \quad i \in S^k, \quad k = L, H \tag{12}$$

where $\hat{\beta}$ or $\hat{\sigma}$ is the winsorized factor.

Notice that the weights in the high factor basket S^H have already been set negative. Thus the final step is to rescale the two baskets according to (8) but keeping the sign in S^H.

The rank weighted method is to first rank the factors of all the stocks and take the weights to be

$$w^i = \bar{z} - z^i \tag{13}$$

where z^i is the rank of stock i and $\bar{z}$ is the average of all the ranks, i.e. $(size(z) + 1)/2$. Then the weighted is rescaled by the same methods as the factor weighted method.

Following one of the two methods above, one could construct the target portfolio. However, in reality, to make the strategy tradable, the portfolio should satisfy several constraints. Normal constraints include turnover, max and min holding positions, max trading liquidity, etc. Here, for our purpose of understanding beta neutrality, we focus on the constraint concerning market neutrality.

We compare two kinds of neutrality, the universal beta neutrality and the sectoral beta neutrality. The final portfolio being S, the neutralities are defined as following

$$\begin{aligned} &\text{C1: universal beta neutrality} \quad \sum_i w^i \beta^i = 0, \quad i \in S, \\ &\text{C2: sectoral beta neutrality} \quad \sum_i w^i \beta^i = 0, \quad i \in Sector\ k \cap S, \quad \forall k, \end{aligned} \tag{14}$$

Note that C2⇒C1

3.1 Backtest results

In Figure 1 and Figure 2, we list the performance of strategies under sector/universe BN constraints.

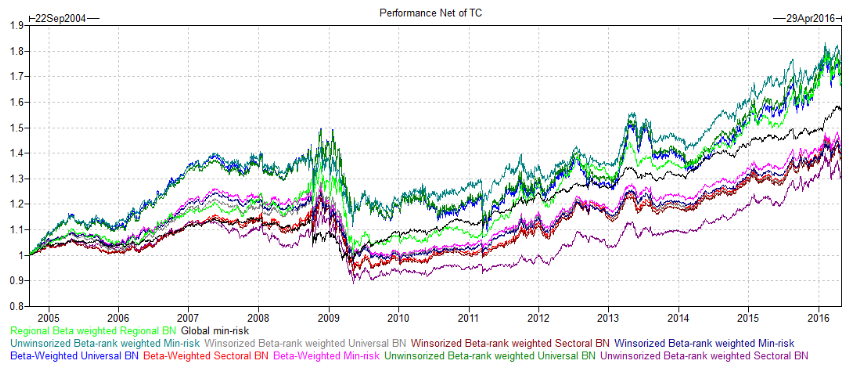

Figure 1: The net of TC performances of low-beta strategies under different conditions

3.2 Conclusions

- requiring region-sector BN has similar results as posing region-sector winsorization. Thus, the results of requiring both is very close to requiring only one of the two.
- using rank or factor value as scores doesn't affect much the performance, it is whether the rank or factor values are winsorized/neutralized region-sectorally or not that greatly affects the performance. One could see clearly in both figures there are two branches, the upper branch are the ones with neither region-sector winsorization nor region-sector Beta neutral, The low branch are the ones with at least one condition.
- Min-risk has similar results as global Beta neutrality on performance.

4 Multi-factor strategies

Now we move from a single factor to the multi-factor world. There are many reasons to consider multi-factor model. The first reason, as described before, is that people tend to look more and more on the factor rather than asset classes or single stocks. They are trying to find anything that could bring excess returns.

More than 350 individual factors have been identified as potential sources of out performance. Below is a review of several single factor products that are

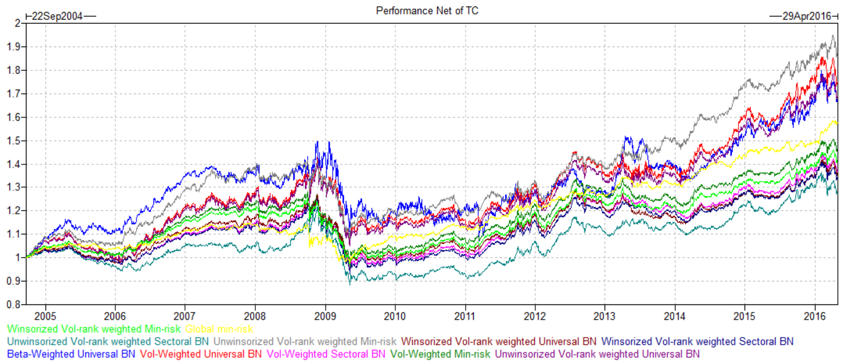

Figure 2: The net of TC performances of low-vol strategies under different conditions

suitable for the construction of multi-factor products. Figure 1 shows factors that have a long history in the academic literature and have been used in the construction of factor indices.

While some have suggested that factor premia are the result of data mining, that if you look hard enough you can get the data to tell you what you'd like to observe, the factors in Table 4 are supported by a wealth of academic literature. In addition, work undertaken by both practitioners and academics has shown that the economic rationale, or the history, behind the factors make good financial sense.

Detractors of the factor investing argument have suggested that these factors might be academically valid, but argue that when they become investment products the crowding or herding effects are likely to lead to an erosion of returns. However, the history of minimum variance investing (the low-volatility factor) suggests that these factors are remarkably robust.

Therefore, if an investor wants to get exposures to more than one factors, it will become really important to know how factors are related and how to combine them in a portfolio.

Another reason that we need a multi-factor model is the risk management. Possessing an accurate estimate of the asset returns covariance matrix is the *sine qua non* of portfolio risk management. How does one calculate such a matrix in practice? The obvious solution is to build a history of asset returns and then calculate the variances and covariances directly. Computing sample statistics directly from historical data, however, is fraught with danger.

Historical returns are typically noisy; even in the absence of actual data

	Explanation	Examples
Value	Undervalued relative to corporate fundamentals	Price-to-book, price-to-earnings
Growth	Above-average earnings growth	Price-to-earnings
Momentum	Rate of acceleration of price	3-month, 6-month, 12-month
Volatility	The dispersion of returns	Volatility VIX
Size	High or low market capitalization	Market cap
Liquidity	Low trading volume	ADV
Quality	Sustainable profitability	Profitability, margins

Table 1: The most common factors

errors, false signals and spurious relationships abound. Two assets may appear closely related when their seemingly-correlated behavior is in fact an artifact of data-mining.

Weak signals and noise aside, when a new asset enters the existing universe, there is no reliable way of calculating its relationships with the other assets, because it does not yet possess a returns history. One could construct various proxies, but such an approach is dubious at best.

Finally, data points totalling no less than the number of assets are required to accurately estimate all the variances and covariances directly. For any realistic number of assets, it is extremely unlikely that sufficient observations exist. Even with a universe of 100 assets, over 5000 relationships need to be estimated. For stock markets like the U.S. (over 12,000 assets), this becomes completely infeasible.

Any one of the above problems is sufficient reason against constructing an asset returns covariance matrix directly. A better approach is to first impose some structure on the asset returns by identifying common factors within the market – that is, factors which drive asset returns. Returns can then be modeled as a function of a relatively small number of parameters, and estimating thousands, or tens, even hundreds of thousands, of asset variances and covariances can thus be simplified to calculating a much smaller handful of numbers.

Generally speaking, factors used in multi-factor models can fall into several broad categories:

- **Fundamental factors**

- **Industry and country factors** reflect a company's line of business and country of domicile.
- **Style factors** encapsulate the financial characteristics of an asset — a company's size, debt levels, liquidity, etc. They are usually calculated from a mixture of market and fundamental (i.e. balance sheet) data.
- **Currency factors** represent the interplay between local currencies of the various assets within the model.
- **Macroeconomic factors** capture an asset's sensitivity to variables such as GNP growth, bond yields, inflation, etc.

- **Statistical factors** are mathematical constructs responsible for the observed correlations in asset returns. They are not directly connected to any observable real-world phenomena, and may change from one period to the next.

Here we will focus on the style factors.

4.1 Multi-factor model

In our analysis, we take the most common five factors (including the BAB factors in Section 2) : Momentum, Value, Quality, Size and BAB.

The multi-factor model says that the return of a stock i at time t is driven by the return of the market and the returns from these factors.

$$r_t^i = \beta_t^i r_t^m + \sum_k x_t^{i,k} \tilde{r}_t^k + \epsilon_t^i \tag{15}$$

where β_t^i is the beta of stock i to the market, $x_t^{i,k}$ is the exposure of stock i to factor k at time t and $\tilde{r}_t^k$ is the return of factor k at time t.

Written in matrix format, it becomes

$$r = \begin{bmatrix} \beta & X \end{bmatrix} \begin{bmatrix} r^m \\ \tilde{r} \end{bmatrix} + \epsilon = \hat{X}\hat{r} + \epsilon \tag{16}$$

where $\hat{X} = \begin{bmatrix} \beta & X \end{bmatrix}$ and $\hat{r} = \begin{bmatrix} r^m \\ \tilde{r} \end{bmatrix}$.

So now we want to estimate the variables in the above equations from the real data. There are two steps. The first step is to estimate $\hat{r}$ which is the factor return. The second step is to estimate the covariance between $\hat{r}$.

There are many possible solutions to this system of equations. If factor exposures X are known, $\hat{r}$ can be estimated using cross-sectional regression analysis. With macroeconomic factors, however, $\hat{r}$ is observed, and it is X rather, that needs to be estimated, typically via time-series regression for each asset. In the case of statistical factors, neither X or $\hat{r}$ is specified, so a rotational indeterminacy exists and both parameters are determined simultaneously, albeit only up to a nonsingular transformation. For a more thorough discussion of multi-factor models and the APT, the curious reader is encouraged to consult [8].

Here we only have style factors, so X is known and we want to estimate $\hat{r}$.

4.1.1 Least square solution

The ordinary least-squares (OLS) regression solution to the factor model of returns seeks to minimize the sum of squared residuals:

$$\hat{r} = \arg\min \sum_i {\epsilon^i}^2 \tag{17}$$

whose solution is straightforward

$$\hat{r} = (\hat{X}^T \hat{X})^{-1} \hat{X}^T r \tag{18}$$

However, the above solution is true only when the following assumptions for OLS hold

1. $\hat{X}$ is a $N \times S$ matrix with full column rank $\rho(\hat{X}) = S$. The OLS solution requires that $\hat{X}^T \hat{X}$ be invertible, which is satisfied only if the columns of $\hat{X}$ are linearly independent. Intuitively, this means the factors should all be distinct from one another.

2. Residuals are zero-mean and independent of the factor exposures . In order for the regression estimates to be unbiased, $E[\epsilon] = 0$ and $E[\hat{X}^T \epsilon] = 0$ are required.

3. Residuals are homoskedastic and have no autocorrelation. These constitute the Gauss-Markov conditions: $Var(\epsilon^i) = \sigma^2$ and $cov(\epsilon^i, \epsilon^j) = 0$ for all $i \neq j$ and establishes the superiority of the least-squares solution over all other linear estimators. Unfortunately, large assets tend to exhibit lower volatility than smaller ones, and homoskedastic residual returns are rarely observed. Figure **??** shows the typical relationship between asset size and returns behavior.

4. Residuals are normally-distributed: strengthening the previous assumption to $\epsilon \sim N(0, \Omega)$ where $\Omega = \sigma^2 I_n$ is not strictly required. Nevertheless, it is a convenient assumption for testing the estimators, to simplify constructing confidence intervals, evaluating hypothesis tests, and so forth.

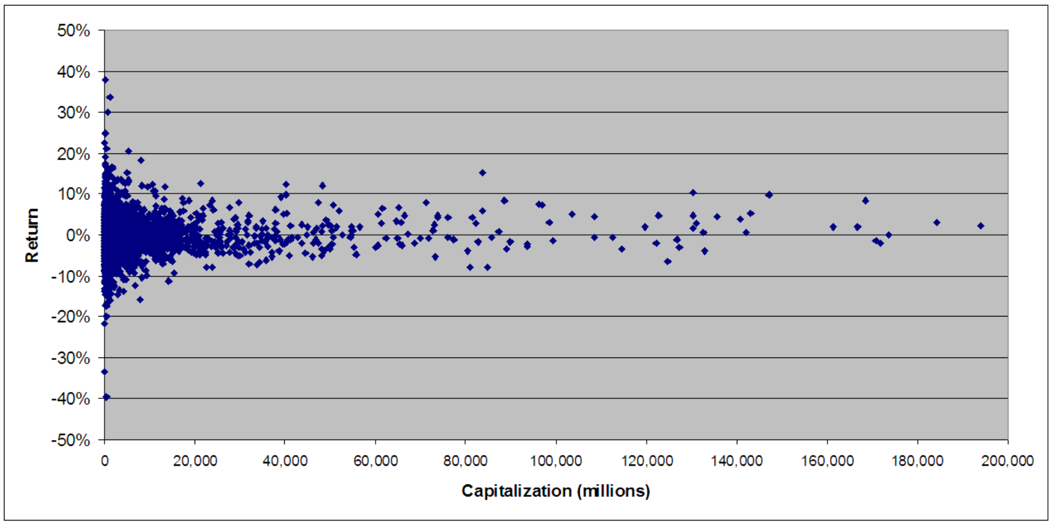

Figure 3: The relation between residual and capitalization

Then how to solve the problem of heteroskedasticity? Traditionally, one corrects for this phenomenon by scaling each asset's residual by the inverse of its residual variance, transforming the above into a weighted least-squares (WLS) problem:

$$W^{1/2} r = W^{1/2} \hat{X} \hat{r} + W^{1/2} \epsilon \tag{19}$$

The solution becomes

$$\hat{r} = (\hat{X}^T W \hat{X})^{-1} \hat{X}^T W r \tag{20}$$

The challenge lies in estimating the residual variances. One could calculate these directly from historical data, but such estimations are noisy and require sufficient history for each asset. As a proxy for the inverse residual variance, one could use the square-root of each asset's market capitalization.

However, when adding other types of factors, especially industry and country factors, there will be the problem of collinearity. For a treatement of such problems, please refer to Appendix **??**.

4.1.2 The risk model

Thus far the discussion has focused entirely on modeling returns, having said nothing whatsoever about the generation of risk forecasts. This is justified — if returns are modeled correctly and robustly then deriving risk estimates is relatively straightforward. For the mathematically-inclined, a more rigorous treatment of the subtleties can be found in [9, 10]. If the model has been sensibly constructed with no important factors missed, then the specific returns are uncorrelated with themselves and with the factors, and the factor risk model can then be derived thus:

$$Q = X\Sigma X^T + \Delta^2 \tag{21}$$

The asset returns covariance matrix Q is a combination of a common factor returns covariance matrix Σ and a diagonal specific variance matrix Δ^2.

The factor covariance matrix is calculated directly from the time-series of factor returns. Regression models estimate a set of factor and specific returns at each time period, eventually building up a returns history.

Recent events should exert more influence on the model than those long in the past, but one cannot simply curtail the history of returns and use only the most recent observations. A sufficiently long history is required to estimate all the covariances reliably. Risk models address this dilemma by weighting the returns matrix using an exponential weighting scheme:

$$w_t = 2^{\frac{t-T}{\lambda}}, t < T \tag{22}$$

where T is the most recent time period. λ is the half-life parameter, the value of t at which the weight is half that of the most recent observation.

The factor covariance matrix is simply calculated as

$$\Sigma = \frac{\hat{r} W \hat{r}^T}{T-1} \tag{23}$$

Selecting an appropriate half-life is a major design question to which there is no definitive answer. This half-life indirectly affects the forecast horizon of the risk model. Too short a short half-life may allow for a very responsive model but creates excessive turnover for asset managers; too long a half-life and the model will fail to respond sufficiently to changing market conditions. The half-life parameter therefore represents a balance between responsiveness and stability. Here our principle is to make the half-life of covariance matrix and the decaying half-life of the signal in the same order.

After having the return model and the risk model, we are now able to construct multi-factor portfolios

4.2 Portfolio construction

A single factor strategy is a strategy that aims at getting exposures to one single factor. The low-beta strategy discussed in Section 2 is an example.

But what should the portfolio be like if one want to get exposures to several factors?

An easy way is to take the average of several single factor strategy, i.e.

$$w_t = \frac{1}{S}\sum_k w_t^k \tag{24}$$

where w_t^k is the portfolio weight of the single factor portfolio at time t.

But the problem is that we are treating every factor equally. So a better idea is to create a new blend target weight and do the optimization based on this target weight.

Here we briefly discuss the method of Generalized Risk Parity as a simple example.

The idea is that in our target portfolio, we want that the risk from each factor is the same. To achieve this, one needs to first define the factor mimicking portfolio.

Starting from (16), one writes it in a cleaner way

$$\hat{r} = (\hat{X}^T W \hat{X})^{-1} \hat{X}^T W r = Mr \tag{25}$$

Therefore, for each factor k (including the market) one has

$$\hat{r}^k = \sum_i M^{ki} r^i \tag{26}$$

Then the k-th row of the matrix M is a portfolio whose return is equal to the return of factor k, we call it the mimicking portfolio of factor k.

Since we already have the risk model, its diagonal terms are then the volatility of the factors.

$$\hat{\sigma} = Diag(\Sigma) \tag{27}$$

One could then derive the risk parity portfolio to be

$$w = M^T \Sigma^{-1} \hat{\sigma} \tag{28}$$

The performance of the portfolio is as Figure 4

Here we tested three versions of GRP strategies, rebalancing daily, every 8 business days, and every 20 business days. But there is no solid reason that the outperformance of 8 days rebalance is not a random effect.

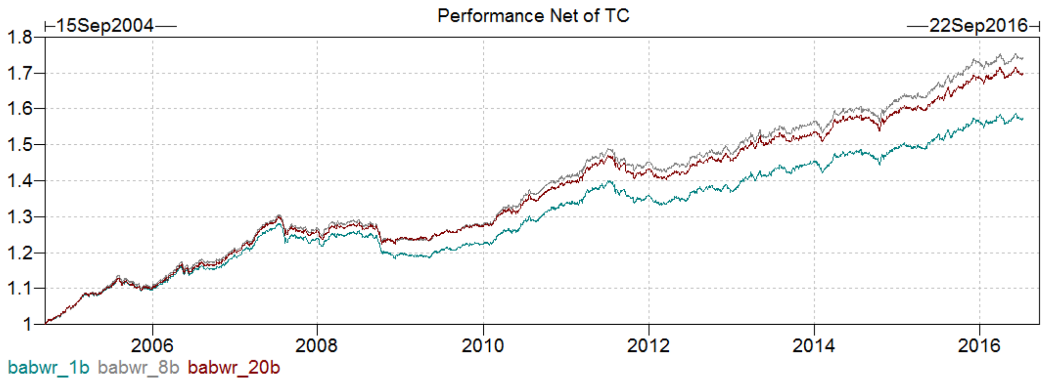

Figure 4: The factor attributions of multi-factor strategy

4.3 Attribution analysis

In the context of multi-factor model, the idea of attribution analysis is to attribute the performance of a strategy to this factors. By this method, one is then able to identify the main driven factor of the performance, whether there is any unexpected factor attribution, etc.

The idea is simple. Starting again from the multi-factor equation (16), for a given portfolio w, one has

$$r = \sum_i w^i r^i = \sum_{i,k} w^i \hat{X}^{ik} \hat{r}^k = \sum_k \left(\sum_i w^k \hat{X}^{ik} \right) \hat{r}^k = \sum_k e^k \hat{r}^k \tag{29}$$

where e^k is the exposure of the portfolio to factor k, and $e^k \hat{r}^k$ is the attribution from factor k.

For a single factor strategy, one should expect that the largest attribution comes from that single factor. For multi-factor strategy, the attribution depends on the method of construction of multi-factor portfolios.

Below is the result of the single factor strategy and multi-factor strategy.

Conclusions: Both the low-beta and the low-vol strategies are mainly driven by BAB factors. And the multi-factor strategy is driven by nearly all the factors. This shows that (1) the BAB factor and low-vol factor are very similar; (2) our method of portfolio construction is able to catch the risk premium from the factor.

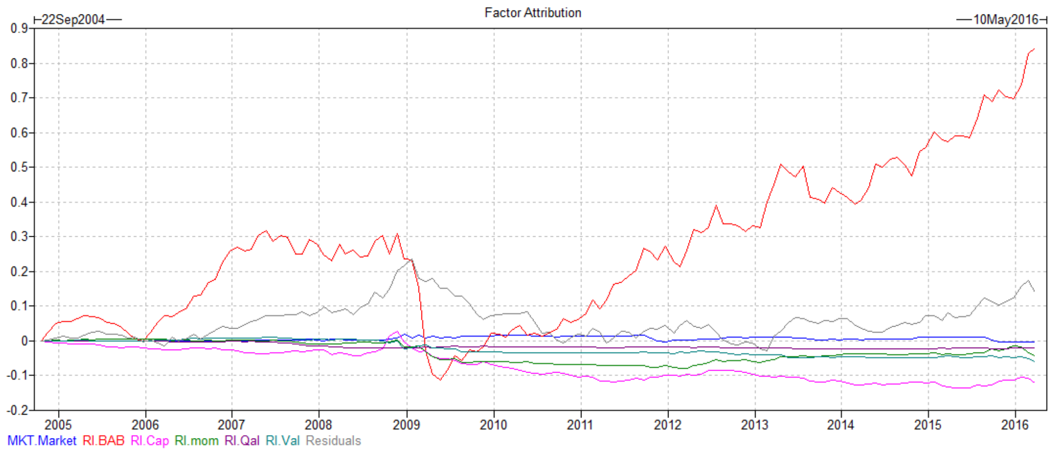

Figure 5: The factor attributions of low-beta strategy

5 The role of constraints in the optimization process

In this final section, we will briefly discuss the role of constraints in the optimization process.

The practical issues that arise due to the interaction between three principal players in any quantitative strategy, namely, the alpha model, the risk model and the constraints are collectively referred to as Factor Alignment Problems (FAP). Examples of FAP include risk underestimation of optimized portfolios, undesirable exposures to factors with hidden and unaccounted systematic risk, consistent failure in achieving ex-ante performance targets, and inability to harvest high quality alphas into above-average IR.

Despite several studies [11, 12, 13, 14, 15, 16], there is considerable disparity in understanding the sources of FAP. While the role of misaligned alpha factors is relatively easy to understand, incorporating the impact of constraints entails considerable analytical complexity that most consultants and researchers find difficult to fathom. A few of them have even gone to the extent of suggesting that aligning alpha and risk factors should suffice in handling FAP. We provide a solid rebuttal to this line of thinking by demonstrating typical symptoms of FAP in optimal portfolios generated by using completely aligned alpha and risk models. Additionally, we provide theoretical guidance to clarify the role of constraints in influencing FAP and illustrate how the Alpha Alignment Fac-

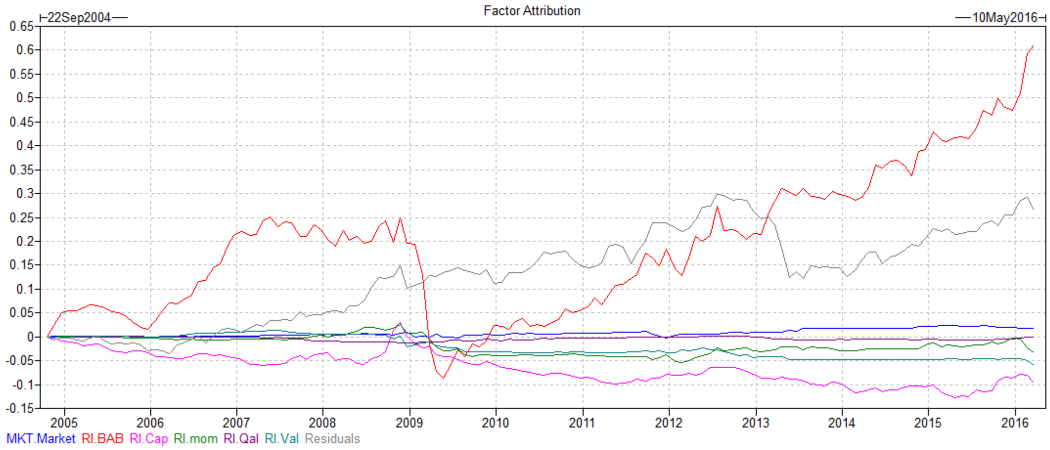

Figure 6: The factor attributions of low-vol strategy

tor (AAF) methodology can handle misalignment resulting from constraints, analytical complexities notwithstanding.

Let's consider a simple example. For the sake of analytical accessibility, we limit our discussion to constrained mean-variance optimization (MVO) problems with a single constraint. All of these results can be easily generalized to encompass several additional constraints.

Consider the following constrained MVO problem with a single factor exposure constraint,

$$\max \alpha^T h - \frac{\lambda}{2} h^T Q h, \quad s.t. \quad \beta^T h \geq \beta_0 \tag{30}$$

We denote this problem $MVO(\alpha, \beta)$.

Let $h(\alpha, \beta)$ denote the optimal solution to $MVO(\alpha, \beta)$. Our goal is to understand the role of the constraint $\beta^T h \geq \beta_0$ in influencing the composition of the optimal portfolio $h(\alpha, \beta)$. In order to pursue this goal, we define two auxiliary unconstrained MVO problems, namely

$$\max \alpha^T h - \frac{\lambda}{2} h^T Q h \qquad MVO(\alpha) \tag{31}$$

$$\max \beta^T h - \frac{\lambda}{2} h^T Q h \qquad MVO(\beta) \tag{32}$$

Let $h(\alpha)$ and $h(\beta)$ denote the optimal solutions to $MVO(\alpha)$ and $MVO(\beta)$, respectively. Note that if $\beta^T h(\alpha) \geq \beta_0$ then $h(\alpha)$ is also an optimal solution

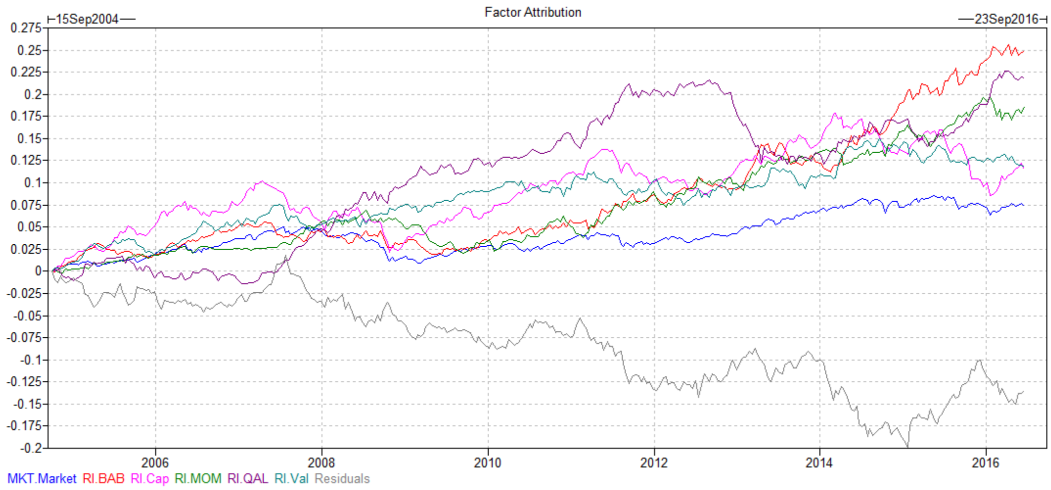

Figure 7: The factor attributions of multi-factor strategy

to $MVO(\alpha, \beta)$ thereby rendering the constraint $\beta^T h(\alpha) \geq \beta_0$ irrelevant. Thus for the purpose of our discussion we assume that $h(\alpha)$ violates the constraint $\beta^T h(\alpha) \geq \beta_0$, and let $\eta = \beta_0 - \beta^T h(\alpha)$ denote the associated constraint violation. Furthermore, without loss of generality we can assume that $\beta^T Q \beta = 1$. The theorem that follows establishes an important link between $h()$, $h()$ and $h(\alpha, \beta)$.

Theorem 1. $h(\alpha, \beta) = h(\alpha) + (\eta\lambda)h(\beta)$.

Theorem 1 shows that the optimal solution to $MVO(\alpha, \beta)$ is obtained by tilting the optimal solution $h(\alpha)$ to the unconstrained problem $MVO(\alpha)$ in the direction $h(\beta)$. Furthermore, the extent of tilting is jointly determined by the risk aversion parameter λ in $MVO(\alpha, \beta)$ and the violation η of the constraint $\beta^T h \geq \beta_0$ by $h(\alpha)$. The higher the risk aversion parameter λ, more significant is the influence of the constraint $\beta^T h \geq \beta_0$ in determining $h(\alpha, \beta)$. Similarly, tighter constraints give rise to higher violation η and consequently have greater influence in determining $h(\alpha, \beta)$.

Theorem 1 also provides additional insights from an alignment perspective. Note that the relationship expressed in Theorem 1 naturally extends to the orthogonal components of $h(\alpha)$, $h(\beta)$ and $h(\alpha, \beta)$. It has been well documented in the literature [13, 15] that optimal portfolios associated with unconstrained MVO problems load up on the orthogonal component of the expected returns. For instance, if $\alpha_\perp \neq 0(\beta_\perp \neq 0)$ then $h(\alpha)(h(\beta))$ will have disproportionately higher exposure to $\alpha_\perp(\beta_\perp)$.

Theorem 1 extends these findings to constrained MVO problems with an intriguing twist. It shows that $h(\alpha, \beta)$ loads up not only on the orthogonal component of α, by virtue of the term $h(\alpha)$, but also on the orthogonal component of β due to the presence of the term $(\eta\lambda)h(\beta)$.

Furthermore, the extent of overloading depends directly on the magnitudes of λ and η. Specifically, highly risk averse strategies that use a higher value of λ, or equivalently lower value of risk targets , are more likely to suffer from misalignment arising from constraints. Of course, if the value of $\lambda(\sigma)$ is very large (small) then the role of constraints diminishes and the portfolio holdings start to resemble minimum variance portfolios, or the benchmark holdings in the case of active strategies.

To summarize, the downward bias in risk prediction that arises exclusively due to the presence of constraints should have a humped shape attaining highest values at moderate risk target levels. By similar arguments, it follows that strategies with tighter constraints leading to higher values of the violation parameter (η) would betray similar characteristics. Until now we have examined results that corroborate the role of constraints in the construction of optimal portfolios. Next we present an interesting result that reverses the roles of alphas and constraints altogether. Consider the following MVO problem.

$$\max \beta^T h - \frac{1}{2\eta} h^T Q h, \quad s.t. \quad \alpha^T h \geq \alpha_0, \quad MVO(\beta, \alpha) \tag{33}$$

Theorem 2 $h(\alpha, \beta) = h(\beta, \alpha)$.

Theorem 2 shows that $MVO(\alpha, \beta)$ and $MVO(\beta, \alpha)$ have identical optimal solutions. In other words, there is nothing sacrosanct about alphas in a constrained MVO problem, and the same optimal portfolio can be obtained by switching the role of alphas and constraints. As an immediate corollary, it follows that the misalignment between constraints and risk factors can have as much influence, if not more, in determining the composition of optimal holdings as that between alpha and risk factors. Furthermore, the relative significance of misalignment due to alpha and constraints can be gauged by comparing the risk-aversion parameters in $MVO(\alpha, \beta)$ and $MVO(\beta, \alpha)$. Specifically, the higher the violation η, the smaller is the risk aversion parameter in $MVO(\beta, \alpha)$ and more prominent is the role of constraints. Notably, the ratio of the risk aversion parameters, namely $\eta\lambda$, is precisely the amount by which $h(\alpha)$ is tilted towards $h(\beta)$ to determine the optimal solution to $MVO(\alpha, \beta)$ (see Theorem 1).

Next we briefly discuss a solution approach, namely the Alpha Alignment Factor (AAF) methodology, to address misalignment arising from constraints. We limit our discussion to key insights and refer the readers to [15] for further details. Since the focus of this paper is misalignment arising exclusively from constraints, we assume that $\alpha_\perp = 0$ in the discussion that follows. Recall that if $\alpha_\perp = 0$, then the only source of misalignment is the orthogonal component of β. In fact, in this case it can be easily shown that the orthogonal component of implied alpha (γ) and $\beta_\perp$ point in the same direction i.e. $\frac{1}{\|\gamma_\perp\|}\gamma_\perp = \frac{1}{\|\beta_\perp\|}\beta_\perp$. The AAF approach recognizes the possibility of systematic risk in the orthogonal component of implied alpha, and penalizes the exposure of the portfolio to $\gamma_\perp$. In our special setting, the AAF optimization problem can be stated as,

$$\max \alpha^T h - \frac{\lambda}{2}(h^T Q h + \nu(h^T y)^2) \quad s.t. \quad \beta^T h \geq \beta_0, \quad (MVO(AAF)) \qquad (34)$$

where $y = \frac{1}{\|\beta_\perp\|}\beta_\perp$, and ν is the systematic risk associated with y. Note that MVO(AAF) can be obtained from $MVO(\alpha, \beta)$ by replacing the covariance matrix Q by an augmented covariance matrix $Q_y = QQ + \nu y y^T$ that has an additional variance term $\nu y y^T$ to capture systematic risks in portfolios by virtue of exposure to $\beta_\perp$. Net we discuss some important characteristics of the optimal solution, say h_y, to $MVO(AAF)$, and compare them with those of $h(\alpha, \beta)$.

Under certain assumptions as laid out in [15], it can be shown that the predicted risk of h_y, namely $\sqrt{h_y^T Q_y h_y}$, is an unbiased estimate of the realized risk of h_y. In other words, while solving $MVO(AAF)$ the optimizer uses an unbiased risk estimate while choosing the optimal portfolio. The same cannot be said about $MVO(\alpha)$. Since the systematic risk of $h(\alpha, \beta)$ that arises by virtue of exposure to $\beta_\perp$ is not captured by Q, and hence goes unaccounted during the optimization phase, it follows that the optimizers ability to select portfolios that have optimal ex-post risk adjusted performance is severely curtailed while solving $MVO(\alpha, \beta)$. This statement can be made precise by using the concept of utility function as described below (see [15] for further details).

Let $U(h) = \alpha^T h \frac{\lambda}{2}\sigma^2(h)$ denote the utility function associated with an arbitrary portfolio h; $\sigma(h)$ denotes the realized risk of h. It can be shown that $U(h_y) \geq U(h(\alpha, \beta))$, and the inequality is strict provided $\beta_X \neq 0$ and $\nu > 0$. Thus using the AAF approach not only gives unbiased risk estimates but also improves the ex-post utility function. Phrased using the concept of efficient frontiers, AAF approach pushes the ex-post frontier upwards thereby allowing the PM to access portfolios that lie above the traditional efficient frontier. The section that follows illustrates this pushing frontier phenomenon using the USER model. To summarize, misalignment arising from constraints is as important and harmful as that arising from misaligned alpha factors. It not only creates statistically significant biases in risk prediction but also obfuscates the ability of the optimizer to solve the quintessential asset allocation problem. AAF approach attacks this problem at its very core; it recognizes the existence of latent systematic risk factors, creates disincentives for the optimizer to load up on such factors, and delivers portfolios that not only have readily available unbiased risk estimates but also superior ex-post risk-adjusted performance.

We conclude this section on an important practical note. Admittedly, the violation parameter η plays a very important role in the narrative presented above. We would like to remind the readers that the violation of constraints by optimal portfolios derived using the unconstrained MVO model is a very common phenomenon; such portfolios are often un-investable due to concentrated long/short positions in certain stocks, excessive turnover, violation of IPS mandates, unacceptable exposures to certain industries/sectors, or simply because they defy common wisdom. Thus constraints are an inextricable component of any quantitative strategy, and as illustrated by the results presented in this paper their contribution to FAP cannot be relegated to secondary considerations.

6 Conclusion

In this report, we first review the economic rationale of the BAB factor, describe the practical way of constructing single factor basket. We then introduce the multifactor model and discussed the estimation methods. We also include here a brief introduction of constructing multi-factor portfolios. Finally, we present a theoretical model on estimating the effects of constraints on the portfolio

Appendices

A Constrained Regression

Our problem is as follows: calculate

$$\min \|Ax - b\|_2, s.t. \quad Bx = d \tag{35}$$

where $A \in R^{n\times k}$,$B \in R^{p\times k}$. We assume the following

- $n \geq k \geq p$
- B has rank p
- A has rank q where $p + q \geq k$

We simultaneously decompose A and B via the generalized singular value decomposition, thus

$$A = U\Sigma X^{-1} \tag{36}$$

$$B = V\Delta X^{-1} \tag{37}$$

and partition the above as follows

$$A = \begin{bmatrix} U_p & U_{k-p} & U_{n-k} \end{bmatrix} \begin{bmatrix} \Sigma_p & 0 \\ 0 & \Sigma_{k-p} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_p^{-1} \\ X_{k-p}^{-1} \end{bmatrix} \tag{38}$$

$$B = V \begin{bmatrix} \Delta_p & 0 \end{bmatrix} \begin{bmatrix} X_p^{-1} \\ X_{k-p}^{-1} \end{bmatrix} \tag{39}$$

We assume without loss of generality that the singular values of A have been sorted so that all zero singular values lie within the block Σ_p. If we make the simple transformation of variables

$$X^{-1}x = y = \begin{bmatrix} y_p \\ y_{k-p} \end{bmatrix} \tag{40}$$

then we may rephrase the problem as

$$\min \|U_p\Sigma_p y_p + U_{k-p}\Sigma_{k-p}y_{k-p} - b\|_2 \tag{41}$$

such that

$$V\Delta_p y_p = d \tag{42}$$

Note that the constraint is now uniquely solvable, viz.

$$y_p = \Delta_p^{-1}V^T d \tag{43}$$

This may be substituted into the minimization equation, resulting in the unconstrained problem

$$\min \|U_{k-p}\Sigma_{k-p}y_{k-p} - \hat{b}\|_2 \tag{44}$$

where

$$\hat{b} = b - U_p\Sigma_p\Delta_p^{-1}V^T d \tag{45}$$

This is easily shown to have the solution

$$y_{k-p} = \Sigma_{k-p}^{-1} U_{k-p}^T \hat{b} \tag{46}$$

And so, we arrive at the final solution

$$y = \begin{bmatrix} \Delta_p^{-1} V^T d \\ \Sigma_{k-p}^{-1} U_{k-p}^T (b - U_p \Sigma_p \Delta_p^{-1} V^T d) \end{bmatrix} \tag{47}$$

References

[1] Ang, Goetzmann and Schaefer *Evaluation of Active Management of the Norwegian Government Pension Fund Global* 2009

[2] Chow et al *A Study of Low Volatility Portfolio Construction Methods* 2013

[3] Black, Jensen and Scholes *The Capital Asset Pricing Model: Some Empirical Test* 1972

[4] Black *Beta and return* 1993

[5] Brennan *Capital market equilibrium with divergent borrowing and lending rates* 1971

[6] Mehrling *Fischer Black and the Revolutionary Idea of Finance* 2005

[7] Frazzini and Pedersen *Betting against beta* 2014

[8] Campbell et al *The Econometrics of Financial Markets* 1997

[9] Zangari *Equity factor risk models. In Modern Investment Management: An Equilibrium Approach* 2003

[10] De Santis et al *Covariance matrix estimation. In Modern Investment Management: An Equilibrium Approach* 2003

[11] Alford et al *Equity portfolio management. In Modern Investment Management: An Equilibrium Approach* 2003

[12] Renshaw et al *Axioma alpha factor method: Improving risk estimation by reducing risk model portfolio selection bias* 2006

[13] Lee and Stefek *Do risk factors eat alphas?* 2008

[14] Saxena and Stubbs *Factor alignment problems and quantitative portfolio management* 2012

[15] Saxena and Stubbs *Alpha alignment factor: A solution to the underestimation of risk for optimized active portfolios.* 2010

[16] Ceria et al *Factor alignment problems and quantitative portfolio management* 2012